

## DICIONÁRIOS E ESTUDOS LINGUÍSTICOS DO PORTUGUÊS MEDIEVAL – UM PROJECTO DE EQUIPA –

---

*Maria Francisca Xavier*

O tema destes Encontros Interdisciplinares da FCSH-UNL – A Ciência na Universidade – sugere que seja feito o enquadramento, embora muito resumidamente, da investigação que vou apresentar na unidade de investigação em que vem sendo desenvolvida – o Centro de Linguística da Universidade Nova de Lisboa – CLUNL.

A entrada da Linguística na Universidade Nova tem uma história quase tão longa como a própria instituição, e o actual Centro de Linguística resulta da reestruturação recente do Centro de Estudos Comparados de Línguas e Literaturas Modernas, criado por decisão do Senado da Universidade com o apoio da então Junta Nacional de Investigação Científica e Tecnológica – JNICT, pouco tempo após a criação, em 1977, da Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa.

No seu início, pretendia-se que o Centro reunisse a investigação literária e linguística dos docentes do Departamento de Línguas e Literaturas Modernas existente na altura, o qual reunia as línguas românicas e germânicas, as literaturas e a linguística. Porém, do mesmo modo que o Departamento se dividiu em cinco, do Centro foram saindo vários grupos ao longo dos anos, até que, há apenas um ano, se concretizou a separação do grupo de Linguística do grupo que restava das Literaturas, dando origem ao Centro de Linguística e a dois outros Centros de Literatura, um de Estudos Anglo-Portugueses e o outro de Estudos Portugueses.

O potencial científico do Centro e a sua capacidade para integrar estudantes na investigação em curso têm, no entanto, dificuldades em se expandir satisfatoriamente devido às sérias limitações de espaço da Faculdade, para as quais o projecto de um Laboratório Associado em instalações pró-

prias é agora a única ideia confortadora, ainda que a sua concretização seja problemática e longínqua, talvez mesmo inviável. A certeza, porém, de que a investigação linguística tem de se projectar em associação com outras áreas, em particular, das ciências sociais e humanas e da engenharia da linguagem, aumenta e assegura o seu interesse e utilidade na formação universitária, bem como na preparação dos estudantes para diversas saídas profissionais.

O orçamento do Centro, até há poucos anos muito reduzido, tem vindo a aumentar nos últimos anos através de financiamentos plurianuais que resultam de candidaturas e das respectivas avaliações externas de quatro em quatro anos. O Centro tem também apresentado projectos a concursos para financiamento específico, inicialmente aos programas da JNICT e a outros ao abrigo de Acordos Internacionais, e mais recentemente aos programas Socrates, Leonardo, Língua, Praxis XXI e Sapiens99. A classificação de muito bom obtida pelo CLUNL na avaliação realizada por peritos estrangeiros, em finais de 1999, vem aumentar o orçamento do Centro, agora acrescido de financiamento programático obtido, também, através de candidatura.

O Centro tem investido, particularmente, em apetrechamento informático e na criação de arquivos e dicionários electrónicos essenciais para o novo modo de trabalhar em Linguística. O CLUNL dispõe de um conjunto considerável de *corpora* textuais informatizados, de bases de dados de dicionários, de terminologias e de bibliografias de grande interesse para a investigação linguística, tanto sincrónica como diacrónica, em especial, da língua portuguesa.

O Centro de Linguística da UNL – CLUNL – engloba a investigação em equipa e os projectos individuais da maioria dos docentes do Departamento de Linguística da FCSH, o que tem reflexos na ligação saudável entre a investigação e a docência. Deste modo, o Centro está organizado em cinco linhas de investigação que cobrem também as áreas disciplinares do ensino graduado e pósgraduado da Linguística na Faculdade, nomeadamente ([www.fcsh.unl.pt/clunl/](http://www.fcsh.unl.pt/clunl/)):

- L.I. 1 – Linguística Comparada, coordenada por M. Francisca Xavier
- L.I. 2 – Lexicologia, Lexicografia e Terminologia, coordenada por M. Teresa Lino
- L.I. 3 – Semântica, coordenada por M. Henriqueta Campos
- L.I. 4 – Linguística Histórica, coordenada por M. Teresa Brocardo
- L.I. 5 – Pragmática, coordenada por Adriano Rodrigues

Os interesses específicos de cada uma das linhas de investigação do CLUNL podem ser resumidos transcrevendo apenas o essencial de pequenos excertos retirados dos objectivos científicos incluídos no último Relatório do novo Centro de Linguística da UNL.



A Linha de Investigação 1 – “Linguística Comparada” – desenvolve actividade em dois domínios complementares: (i) estudos linguísticos, no âmbito do léxico e da sintaxe comparada relativamente aos processos de aquisição, variação e mudança, e (ii) criação de *corpora* textuais informatizados, de bases de dados e de dicionários electrónicos, tanto do Latim tardio como do Português Medieval e do Português Língua nativa e não nativa.

A Linha de Investigação 2 – “Lexicologia, Lexicografia e Terminologia” – enquadra a sua investigação em modelos teóricos e metodológicos em constante articulação com a Linguística e a Informática (Informática de Orientação Textual, Lexicomática e Terminótica) e, por vezes, em interdependência com as Novas Tecnologias da Informação. Criou e desenvolve arquivos electrónicos de terminologias e de *corpora* em diversos domínios.

A Linha de Investigação 3 – “Semântica” – procura descrever e explicar factos linguísticos do Português Europeu no quadro da Teoria das Operações Predicativas e Enunciativas. Visa a construção de gramáticas locais a partir de temas específicos nos domínios da determinação nominal e verbal, tendo presente a singularidade de cada termo e a sua interrelação com outros termos na cadeia de operações subjacentes à actividade da linguagem.

A Linha de Investigação 4 – “Linguística Histórica” – tem privilegiado, entre outras, questões que se prendem com o tratamento de fontes para estudos linguísticos diacrónicos, critérios para o estabelecimento e codificação de *corpora*, análise linguística de diversos tipos de fontes textuais, aspectos da morfologia histórica do português e está a desenvolver uma base de dados bibliográficos de História da Língua Portuguesa.

A Linha de Investigação 5 – “Pragmática” – foi recentemente constituída com o objectivo de proceder à análise de um *corpus* de interacções institucionais.

É neste enquadramento institucional que se desenvolvem os “Dicionários e os Estudos Linguísticos do Português Medieval” de que se pretende dar conta aqui.

Trata-se, fundamentalmente, da investigação que uma equipa interdisciplinar tem vindo a realizar, tendo por base edições de textos antigos de reconhecido interesse linguístico, histórico e cultural que constituem o *Corpus* Informatizado do Português Medieval – CIPM – do CLUNL.

O CIPM está a ser construído desde 1993 com a colaboração de linguistas, historiadores, estudiosos de literatura e de cultura medieval da FCSH, da FLUL e da Universidade de Oxford, e também com a presença constante de estudantes. Actualmente com mais de três milhões de palavras, o CIPM incorpora textos latinos da actual região portuguesa dos séculos IX a XII e textos portugueses dos séculos XII a XVI.

O início da constituição do CIPM realizou-se no âmbito de um projecto intitulado “Gramática do Português Medieval. Contributos para a sua



Caracterização”, subsidiado pela ex-JNICT de 1993 a 1996, e o alargamento posterior do CIPM, entre 1996-2000, foi financiado ao abrigo do projecto “*Corpora do Português Medieval. Etiquetagem e Segmentação Automáticas*”, FCT/PRAXIS XXI: 2/2.2/CSH/778/95.

Para dar uma ideia do trabalho desenvolvido com vista ao alargamento do CIPM, transcreve-se a seguir a lista dos textos tratados no ano de 1999/2000. Foram, então, digitalizados e/ou corrigidos, adaptados, anotados e integrados no CIPM os seguintes textos do século XIII ao século XV<sup>1</sup>:

- TOX – *Textos Notariais do Arquivo de Textos do Português Antigo* de Stephen Parkinson (ed.) nova versão em suporte digital (Dezembro de 1998).
- DN – *Documentos Notariais dos Séculos XII a XVI* de A.M. Martins (ed.) nova versão em suporte digital (1999).
- VS – *Vidas de Santos de um Manuscrito Alcobacence* de Ivo Castro *et alii* (eds.) (1985) *Vidas de Santos de um Manuscrito Alcobacence* (Cod. Alc. CCLXVI / ANTT 2274), Lisboa, I.N.I.C. pp. 16-52; 59-83.
- CGE – *Crónica Geral de Espanha de 1344* de Luís Filipe Lindley Cintra (ed.) (1951) *Crónica Geral de Espanha de 1344*, Lisboa, I.N.C.M.
- CAXL/CAXP – *Crónica de Afonso X* (Ms L/P) de Luís Filipe Lindley Cintra (ed.) (1951) *Crónica Geral de Espanha de 1344*, Lisboa, I.N.C.M.
- HRP – *História dos Reis de Portugal* de Luís Filipe Lindley Cintra (ed.) (1951) *Crónica Geral de Espanha de 1344*, Lisboa, I.N.C.M.
- CRB – *Chronica dos Reis de Bisnaga* de David Lopes (ed.) (1897) *Chronica dos Reis de Bisnaga*, Lisboa, Imprensa Nacional.
- LE – *Livro da Enseñança de Bem Cavalgar Toda Sela* de Joseph M. Piel (ed.) (1944) *Livro da Enseñança de Bem Cavalgar Toda Sela*. Lisboa, Bertrand. (Texto editado com emendas posteriores de João Dionísio e separação de palavras da bolseira Ana Castro).
- LC – *Leal Conselheiro* de Joseph M. Piel (ed.) (1942) *Leal Conselheiro*, Lisboa, Livraria Bertrand. (Texto editado com emendas posteriores de João Dionísio e separação de palavras da bolseira Ana Castro).
- CP – *Castelo Perigoso* de J. A. S. Neto (ed.) (1997), *Duas Leituras do Tratado Ascético-Místico Castelo Perigoso*, Dissertação de Doutoramento, São Paulo, Faculdade de Filosofia, Letras e Ciências Humanas, USP. (Texto editado com emendas posteriores de Irene Nunes com a colaboração do bolseiro Manuel Freitas, pp. 125-243).

---

<sup>1</sup> Veja-se a constituição actual do CIPM em Xavier; Crispim (2000) e em <http://cipm.fcsh.unl.pt/>.



- OE – *Orto do Esposo* de B. Maler, (ed.) (1956) *Orto do Esposo. Texto Crítico*, Vol. I, Rio de Janeiro, M.E.C., Instituto Nacional do Livro.
- CD – *Chancelarias Portuguesas: D. Duarte (1433-1435)* de J.A. Dias (ed.) (1998/99) *Chancelarias Portuguesas. D. Duarte*. Vol. I, Tomos 1 e 2; Vol. II. Lisboa, Centro de Estudos Históricos da UNL, com a colaboração do bolseiro Pedro Pinto.

Não existindo um Dicionário do Português Medieval, um dos principais interesses da equipa interessada nos estudos sobre o Português Medieval tem sido a preparação de glossários, de dicionários electrónicos e de listagens de palavras com a respectiva classificação morfológica dos textos inseridos no CIPM, como uma etapa prévia para a elaboração do Dicionário.

Para facilitar a extracção de informação dos *corpora* textuais tem-se procurado o auxílio de ferramentas informáticas. Com o desejo de alcançar este objectivo, constituiu-se, em 1995, uma equipa mista de linguistas e de informáticos, estes últimos da Faculdade de Ciências e Tecnologia da UNL.<sup>2</sup>

A primeira ferramenta a ser desenvolvida e adaptada para analisar os textos do século XIII do CIPM foi precisamente um etiquetador morfossintáctico automático, o qual atribui a cada palavra *corpus* textual uma etiqueta de classe de palavra, facilitando de imediato a identificação das palavras que irão ser objecto de estudo lexicográfico.<sup>3</sup>

A etapa seguinte de preparação do Dicionário consistiu na concepção de uma metodologia que permitisse trabalhar os dados, analisando-os em fases sucessivas relativamente às propriedades morfológicas, sintácticas e semânticas, com vista à sua inclusão nas entradas de Dicionário<sup>4</sup>.

Um primeiro volume de um *Dicionário de Verbos do Século XIII*, organizado por M. Francisca Xavier; Graça Vicente e M. Lourdes Crispim, foi publicado em 1999. Foram também concebidos e iniciados ao abrigo do projecto FCT/PRAXIS XXI: 2/2.2/CSH/778/95 mais dois dicionários: um *Dicionário de Nomes Comuns e de Nomes Próprios*, e ainda um *Dicionário de Termos* que deu origem a uma base de dados para a *Terminologia Portuguesa Histórica – Jurix*<sup>5</sup>.

Relativamente ao *Dicionário de Nomes* foi elaborada uma sub-

---

<sup>2</sup> Veja-se Xavier et alii (1999) e, também, Xavier, Lopes, Marques, Rocio e Silva (no prelo) para uma descrição da utilização de ferramentas informáticas no tratamento de textos do CIPM.

<sup>3</sup> O etiquetador foi concebido por Nuno Marques e Gabriel Pereira Lopes e a sua adaptação ao Português Medieval foi financiada pela JNICT/FCT.

<sup>4</sup> Veja-se Xavier e Vicente (1997) e Xavier (no prelo<sup>a</sup>).

<sup>5</sup> Veja-se Xavier; Castro e Gonçalves (no prelo) e, também em breve, em /www.cipm.fcsh.unl.pt/.

-classificação das classes de Nomes Próprios e de Nomes Comuns que ocorrem em textos não literários do século XIII<sup>6</sup>.

Assim, as subclasses de Nomes Próprios (NP) são:

NP – *Deus, Santa Trindade...*

NPA – Nome próprio de pessoa (antropónimo) Ex.: *dona Marina Perez, São João...*

NPT – Nome próprio de lugar (topónimo) Ex.: *Monte Mayor o Velho, freguesia de Santa Maria Madalena...*

NPC – Nome próprio de instituições, etc. Ex.: *ordem de Avis, concelho de Monsaraz...*

NA – Nome designativo de pessoa específica com propriedades descritivas e que inclui um NP. Ex.: *abadessa de Achelas, mestre de Avis...*

NT – Nome designativo de lugar específico, sendo duvidoso se se trata de nome próprio ou de descrição definida. Ex.: *fonte do álamo, lagoa grande, lagoa segunda...*

ND – Datas: dias, meses, anos, dias festivos... Ex.: *dia de natal, II de agosto da era de MCCXLII, São Miguel de Outubro...*

E as subclasses de Nomes Comuns são:

NC – nomes comuns gerais. Ex.: *água, amor, terra...*

NCp – nomes de profissões/profissionais, cargos/detentores desses cargos, parentesco/parentes. Ex.: *chanceler, mestre, pais, prestameiro, herdeiro, sucessor, povoador.*

NCa – nomes de animais. Ex.: *porco, galinha...*

NCt – nomes topográficos. Ex.: *campo, herdade, fonte, leira, povoação, vereda...*

NCm – nomes metrológicos. Ex.: *moio, spadoa...*

À medida que outros textos vão sendo estudados, novas classes deverão ser incorporadas no *Dicionário de Nomes do Português Medieval*.

A *Terminologia Portuguesa Histórica*, iniciada recentemente, pode ser consultada numa base de dados – *Jurix* – que dispõe dos seguintes campos:

---

<sup>6</sup> A investigação que levou à preparação de glossários e à concepção dos Dicionários de Nomes e de Termos foi financiada pela FCT/PRAXIS XXI.



- (i) Vedeta do termo,
- (ii) Informação gramatical,
- (iii) Pronúncia,
- (iv) Etimologia,
- (v) Fonte da etimologia,
- (vi) Definição,
- (vii) Fonte da definição,
- (viii) Variantes gráficas,
- (ix) Fontes textuais da(s) ocorrência(s),
- (x) Temas em que se integra o termo,
- (xi) Termos relacionados,

Por exemplo:

- (i) Vedeta do termo: *herdamento*
- (ii) Informação gramatical: Nome
- (iii) Pronúncia: –
- (iv) Etimologia: De *herdar*+*-mento*
- (v) Fonte da etimologia: DELP (*Dicionário Etimológico da Língua Portuguesa*)
- (vi) Definição: Herança ou propriedade (herdada)
- (vii) Fonte da definição: *Foro Real*, vol. 2, pág. 151
- (viii) Variantes gráficas: *erdamento*, *e(r)dam(en)to*, *erdam(en)to*, *erdam(en)tos*, *e(r)dame~to*, *erdamentho*, *erda[me~th]o*, *erda-me~tho*, *h(er)dametos*
- (ix) Fontes textuais da(s) ocorrência(s): *Chancelaria de Afonso III* 012; 027; *Documentos Notariais* 003; 019; 040; 041; 042; 095; *História do Galego-Português* 058...
- (x) Temas em que se integra o termo: câmbio, doação, emprazamento, procuração, relato/testemunho, sentença, testamento
- (xi) Termos relacionados: herança, herdade, herdadura, manda, testamento

Com o intuito de consolidar o trabalho em equipa e com objectivos bem definidos a serem atingidos em períodos determinados, dois projectos foram submetidos ao programa de financiamento Sapiens 99, em Janeiro de 2000, e embora apenas um tenha directamente a ver com os estudos sobre o Português Medieval, o segundo é também interessante e útil para estes estudos, porque propõe uma discussão sobre aspectos actuais da reflexão linguística teórica que são, necessariamente, pertinentes para a explicação dos factos de variação e mudança da História da Língua Portuguesa.

- (I) Sapiens99:  
“Léxico e Sintaxe do Verbo no Português Medieval”  
Responsável: Maria Francisca Xavier

O projecto tem como objectivo contribuir para um melhor conhecimento do léxico e do sistema linguístico na origem e na história da língua portuguesa, desde o período pré-histórico até aos finais do século XV, através do estudo da estrutura da frase e da sintaxe do verbo numa perspectiva comparada.

Os dados empíricos são extraídos do CIPM e estão a ser estudados pelo grupo em dois campos intimamente relacionados – o léxico e a sintaxe do Português Medieval.

No que respeita ao léxico, continuará a ser desenvolvido um dicionário de verbos – na linha de Xavier, Vicente e Crispim *orgs.* (1999) *Dicionário de Verbos Portugueses do Século XIII*, Edição da L.I. 1 do CLUNL.

- (i) Constituirão o *corpus* para o estudo das entradas lexicais alguns textos latinos dos séculos XI-XII, os textos portugueses dos séculos XII-XIII recentemente informatizados, e alguns textos dos séculos XIV e XV, todos integrados no CIPM.
- (ii) Todas as formas verbais deverão ser identificadas, extraídas e morfológicamente classificadas, conservando as variantes gráficas e a informação relevante da sua fonte textual.
- (iii) Serão preparadas as concordâncias necessárias.
- (iv) Com base na análise das frases delimitadas nas concordâncias, e recorrendo também aos textos completos, sempre que for necessário, serão identificados os contextos que permitem descrever a informação lexical de cada verbo.
- (v) As entradas lexicais incluem:
  - a) a vedeta do verbo (+verbos relacionados),
  - b) a etimologia,
  - c) as formas gráficas atestadas classificadas,
  - d) o(s) significado(s) em Português Actual,
  - e) a(s) estrutura(s) argumental/is,
  - f) a(s) estrutura(s) de subcategorização,
  - g) o(s) exemplo(s).

A informação trabalhada deverá ser introduzida na base de dados do *Dicionário de Verbos do Português Medieval – DVPM*.<sup>7</sup>

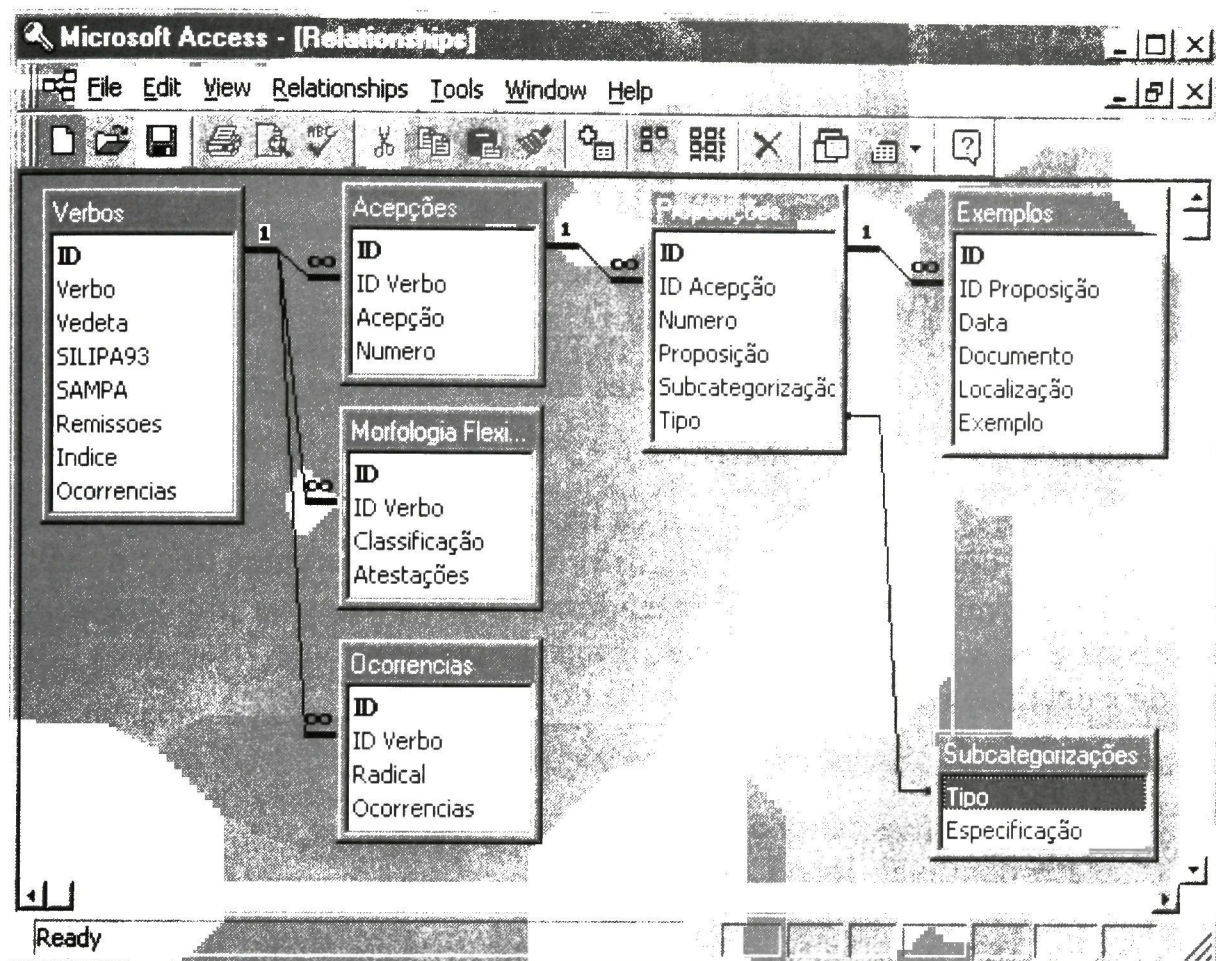
---

<sup>7</sup> Concebida com o apoio informático do Eng. Inf. Luís Reis.



## Base de Dados do DVPM

### Definição de Relações no Access



O trabalho realizado para a preparação da informação sobre os verbos a inserir na base de dados do DVPM é igualmente útil para outros estudos linguísticos. Efectivamente, o estudo da estrutura da frase e da sintaxe do verbo numa perspectiva diacrónica e comparada está a ser desenvolvido através da análise sistemática das construções em que ocorrem os verbos das diferentes classes. Em seguida, tipificam-se as ordens de constituintes pertinentes e procura-se caracterizar as projecções funcionais das estruturas das orações seguintes:

- infinitivas complementos de verbos de controlo, de elevação, modais, causativos e perceptivos – M. Cristina V. da Silva
- transitivas, intransitivas e inacusativas – M. Alexandra Fiéis
- adjuntas não finitas com verbos flexionados e não flexionados – Maria Lobo
- possessivas com os verbos “ter”, “haver”, “possuir”, “ser” e “estar” – Ana Castro

A reflexão continuada e a discussão aberta do grupo com outros investigadores são estimuladas, por um lado, pela focalização num tema abrangente – o estudo da sintaxe do verbo e da estrutura da frase –, por outro lado, pela necessidade de confrontar as diferentes teorias e propostas avançadas na explicação de fenómenos lexicais e morfosintácticos de outras línguas, dialectos e estados de língua, tais como:

- (i) os sistemas de auxiliar;
- (ii) as diversas realizações de caso;
- (iii) as ordens de constituintes diferentes;
- (iv) a inversão do sujeito;
- (v) o posicionamento variável dos pronomes clíticos;
- (vi) a interpolação de constituintes entre o clítico e o verbo;
- (vii) determinantes e pronomes.

Espera-se conseguir identificar os movimentos estritamente sintácticos, separando-os dos artifícios estilísticos próprios dos diferentes tipos de textos, o que permitirá um melhor entendimento dos processos implicados na mudança e na variação linguísticas, tornando possível uma visão mais profunda e detalhada da sintaxe diacrónica do Português.

(II) Sapiens99:

(Non-)Universality of functional categories

Responsável: João Costa

O objectivo do projecto é desenvolver investigação sobre o debate respeitante ao domínio funcional da estrutura da frase.

Duas hipóteses são defendidas na literatura recente:

- (i) as categorias funcionais são projectadas uniformemente e universalmente;
- (ii) poderá haver variação linguística interna a uma língua e variação nas diferentes línguas relativamente ao domínio funcional da frase.

Está em curso o estudo de várias construções defectivas de maneira a poder ser avaliada a validade da hipótese de que as categorias funcionais só são projectadas quando é necessário, nomeadamente:

- construções gerundivas e participiais;
- construções causativas;
- topicalização em contextos não finitos;
- fala telegráfica das crianças.



Serão construídos testes empíricos para determinar se a estrutura funcional é ou não projectada nestes domínios, e serão avançados argumentos a favor ou contra a universalidade das categorias funcionais.

Uma das grandes vantagens da complementariedade de projectos de natureza teórica e descritiva, simultaneamente, de dados sincrónicos e diacrónicos reside no facto de ser possível, relativamente à sincronia, construir testes empíricos que, relativamente à diacronia, não são viáveis, podendo estes contribuir para formular hipóteses de explicação dos fenómenos que os dados históricos revelam.

Os resultados dos projectos serão apresentados em encontros científicos e publicados em livro e na Internet. O interesse demonstrado por investigadores portugueses, brasileiros, espanhóis e italianos pela consulta das bases de dados e dos arquivos electrónicos do Português Medieval leva a que a disponibilização de *corpora* do CIPM e das bases de dados dos Dicionários através da Internet seja um dos objectivos a alcançar em 2001.

O trabalho aqui relatado só tem sido possível porque nele têm colaborado muitas pessoas. Vários colegas e estudantes deixaram aqui as suas contribuições e seguiram os seus caminhos. Com a ajuda de muitos conseguimos organizar, digitalizar e classificar muita informação útil para futuras investigações e estamos em condições de a disponibilizar para todos os interessados em desenvolver estudos sobre o Português Medieval. Só uma equipa periodicamente revista mas sempre aberta a colaborações externas consegue realizar trabalho de natureza interdisciplinar que se revela de interesse para diversos domínios.

## **Equipa e Colaboradores em 2000**

### *Linguística*

M. Francisca Xavier; M. de Lourdes Crispim; Graça Vicente; João Costa; Ana Madeira; Maria Lobo; M. Alexandra Fiéis; M. Cristina V. Silva; Ana Castro

### *Corpora*

J. Alves Dias (C.E.H. da UNL); Helder Godinho e Irene Nunes (I.L.M. da FCSH-UNL); João Dionísio (Românicas da FL-UL); Ana M. Martins (Linguística da FL-UL); Stephen Parkinson (U. Oxford); Pedro Pinto; Manuel Freitas; Mafalda Proença e Ana Cristina Gonçalves (L.I. 1 do CLUNL)

## **Processamento de língua natural**

Gabriel Pereira Lopes; Vitor Rocio; Joaquim F. da Silva e Mário Alves (FCT-UNL); Nuno Marques (U. Aberta/ FCT-UNL)

## Projectos sobre o Português Medieval financiados até 2000

- “Gramática do Português Medieval”, subsidiado em 1993 ao abrigo do Acordo JNICT-British Council. Teve como objectivo projectar os estudos a desenvolver em equipa constituída por docentes da FCSH-UNL, com Stephen Parker da U. de Oxford e Roger Wright da U. de Liverpool.
- “Gramática do Português Medieval. Contributos para a sua Caracterização”, subsidiado pela ex-JNICT, de 1993 a 1996. Este projecto deu início à constituição do CIPM, ao *Dicionário de Verbos Portugueses do Século XIII* e a outros estudos, de que se destaca uma dissertação de Mestrado sobre a sintaxe dos clíticos em textos do século XIII (Fiéis 1997).
- “Etiquetagem e Segmentação Automáticas de *Corpora* de Português Medieval”, subsidiado de 1996 a 1999 pela JNICT/FCT. Foram objectivos deste projecto, por um lado, a constituição de uma equipa mista de linguistas e de informáticos, por outro lado, a adaptação e utilização de ferramentas para classificação e análise morfo-sintáctica em textos portugueses medievais.
- “*Corpora* do Português Medieval. Etiquetagem e Segmentação Automáticas”, subsidiado de 1997 a 2000 pela FCT/PRAXIS XXI: 2/2.2/CSH/778/95. Este projecto permitiu o alargamento cronológico e tipológico do CIPM desenvolvido por uma equipa interdisciplinar constituída por linguistas, historiadores e estudiosos da literatura e da cultura portuguesas medievais, assim como a elaboração de glosários, o estudo de terminologia antiga, o prosseguimento dos estudos linguísticos já anteriormente iniciados bem como a utilização de diversas ferramentas automáticas para extracção, análise e classificação de palavras, expressões e frases (Rocio et alii e Xavier et alii).

## Referências

- Castro, A. (1998) “Sintagmas Nominais e a Subida de N; estudo comparativo entre o Italiano, o Português Europeu Contemporâneo e o Português Antigo”, in *Actas do XIV Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa, APL.
- Fiéis, M. A. (1997)) Fiéis, M. A. (1996) *Clíticos num Corpus do Português do Século XIII*, Diss. de Mestrado, Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa.
- Fiéis, M. A. (no prelo) “Pares Ergativos em Textos Portugueses do Século XIII”, in *Actas do Congresso Internacional “500 Anos da Língua Portuguesa no Brasil”*, Évora, University of Évora.
- Fiéis, M. A. (no prelo) “Estudo Diacrónico da Interpolação em Português”, in *Actas do Congresso Internacional de Linguística “Léxico e Gramática”*, Coimbra, Almedina.
- Rocio, V., M. Alves, G.P. Lopes; M.F. Xavier; M.G. Vicente (2000) “Automated creation of a partial treebank of Medieval Portuguese”, in Anne Abeillé (ed.). *Building and using syntactically annotated corpora*, Dordrecht, Kluwer Academic Publishers (Language and Speech Series).



- Silva, C. V. (no prelo) “Uma análise diacrónica das construções com verbos modais em Português” in *Actas do XVI Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa, APL.
- Silva, C. V. (no prelo) “Uma análise diacrónica das construções causativas em Português”, in *Actas do Congresso Internacional de Linguística “Léxico e Gramática”*, Coimbra, Almedina.
- Silva, C. V. (no prelo) “As construções causativas em textos notarias dos séculos XI a XIV”, in *Actas do Congresso Internacional “500 Anos da Língua Portuguesa no Brasil”*, Évora, Universidade de Évora.
- Xavier, M.F.; M.G. Vicente (1997) “A Problemática de um Dicionário de Verbos do Século XIII” in *Sentido que a Vida Faz. Estudos para Óscar Lopes*, Porto, Campo das Letras, 897-904.
- Xavier, M. F.; M. L. Crispim; G. Vicente; A. Castro; A. Fiéis; C. V. Silva; M. Lobo (1999) “Utilizações Informáticas de *Corpora* Textuais Medievais” in Marrafa, P. & M. A. Mota (orgs.) *Linguística Computacional. Investigação Fundamental e Aplicações*, Lisboa: APL/Colibri, 347-358.
- Xavier, M. F.; G. Vicente e M. L. Crispim orgs. (1999) *Dicionário de Verbos Portugueses do Século XIII*, Edição da L.I. 1 do CLUNL.
- Xavier; M. F. e M. L. Crispim (no prelo) “*Corpus* Informatizado do Português Medieval – CIPM. Constituição e Processamento Automático” in *Anais do 6º. Congresso da Associação Internacional de Lusitanistas*, Rio de Janeiro, 1999.
- Xavier, M. F. (no prelo<sup>a</sup>) “Caso Inerente – Variação e Mudança no Português” in *Actas do Congresso Internacional 500 Anos da Língua Portuguesa no Brasil 2000*, Universidade de Évora.
- Xavier, M. F. (no prelo<sup>b</sup>) “Informação Sintáctica e Semântica num Dicionário de Verbos do Português Medieval” in *Actas do Congresso Internacional ‘Léxico & Gramática’ Lugo 2000*, Coimbra, Almedina.
- Xavier, M. F.; A. Castro; A. C. Gonçalves (no prelo) “A Mais Antiga Terminologia Notarial Portuguesa” in *Actas do Congresso Internacional 500 Anos da Língua Portuguesa no Brasil 2000*, Universidade de Évora.
- Xavier, M. F., G. P. Lopes; N. Marques; V. Rocio; J. F. Silva (no prelo) “Ferramentas Informáticas para o Português Medieval” in *Actas do Congresso Internacional 500 Anos da Língua Portuguesa no Brasil 2000*, Universidade de Évora.